Math 164: Problem Set 5

Owen Jones

2/9/2024

6.29 (a) $\frac{1}{n}\sum_{i=1}^{n}(ax_i + b - y_i)^2$

$= \frac{1}{n}\sum_{i=1}^{n} a^2x_i^2 + b^2 + y_i^2 - 2ax_iy_i + 2ax_ib - 2by_i$

$= a^2\overline{X^2} + b^2 + \overline{Y^2} + 2a\overline{XY} + 2ab\overline{X} - 2b\overline{Y}$

$= \mathbf{z}^\top \begin{bmatrix} \overline{X^2} & \overline{X} \\ \overline{X} & 1 \end{bmatrix} \mathbf{z} - 2\begin{bmatrix} \overline{XY} & \overline{Y} \end{bmatrix} \mathbf{z} + \overline{Y^2}$

(b) By the FONC $\nabla f = 2\begin{bmatrix} \overline{X^2} & \overline{X} \\ \overline{X} & 1 \end{bmatrix} \mathbf{z} - \begin{bmatrix} \overline{XY} \\ \overline{Y} \end{bmatrix} = 0$

$\Rightarrow \mathbf{z}^* = \begin{bmatrix} \overline{X^2} & \overline{X} \\ \overline{X} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \overline{XY} \\ \overline{Y} \end{bmatrix}$ is the only solution.

$= \frac{1}{\overline{X^2} - \overline{X}^2}\begin{bmatrix} 1 & -\overline{X} \\ -\overline{X} & \overline{X^2} \end{bmatrix} \begin{bmatrix} \overline{XY} \\ \overline{Y} \end{bmatrix}$

$= \begin{bmatrix} \frac{\overline{XY} - \overline{X}\cdot\overline{Y}}{\overline{X^2} - \overline{X}^2} \\ \frac{\overline{X^2}\cdot\overline{Y} - \overline{XY}\cdot\overline{X}}{\overline{X^2} - \overline{X}^2} \end{bmatrix}$

(c) WTS $\overline{Y} = \overline{X}\frac{\overline{XY} - \overline{X}\cdot\overline{Y}}{\overline{X^2} - \overline{X}^2} + \frac{\overline{X^2}\cdot\overline{Y} - \overline{XY}\cdot\overline{X}}{\overline{X^2} - \overline{X}^2}$

$= \frac{\overline{X}\cdot\overline{XY} - \overline{X}^2\cdot\overline{Y}}{\overline{X^2} - \overline{X}^2} + \frac{\overline{X^2}\cdot\overline{Y} - \overline{XY}\cdot\overline{X}}{\overline{X^2} - \overline{X}^2}$

$= \frac{\overline{Y}(\overline{X^2} - \overline{X}^2)}{\overline{X^2} - \overline{X}^2}$

$= \overline{Y}$

6.30 Let $f(\mathbf{x}) = \frac{1}{p}\sum_{i=1}^{p}\|\mathbf{x} - \mathbf{x}^{(p)}\|^2 = \frac{1}{p}\sum_{i=1}^{p}(\mathbf{x} - \mathbf{x}^{(p)})^\top(\mathbf{x} - \mathbf{x}^{(p)})$

$\Rightarrow \nabla f = \frac{1}{p}\sum_{i=1}^{p} 2(\mathbf{x} - \mathbf{x}^{(i)}) = 0$

$\Rightarrow \arg\min_{\mathbf{x}} f(\mathbf{x}) = \mathbf{x}^* = \frac{1}{p}\sum_{i=1}^{p}\mathbf{x}^{(i)}$ which is just the mean (or centroid).

Because the hessian $F(\mathbf{x}) = \mathbf{I}_n > 0$, the mean of f is a strict local minimizer.

6.31 Because Ω is convex, let $\phi(\alpha) = f(\mathbf{x}^* + \alpha\mathbf{d})$ for some feasible direction $\mathbf{d}$ and $0 < \alpha_0$ s.t $\mathbf{x}^* + \alpha\mathbf{d} \in \Omega, \forall\alpha \in [0, \alpha_0]$. By the MVT we have $\frac{\phi(\alpha)-\phi(0)}{\alpha} = \mathbf{d}^\top\nabla f(\xi)$ for $0 \leq \xi \leq \alpha$. Because $\lim_{\alpha\to 0} \frac{\phi(\alpha)-\phi(0)}{\alpha} = \mathbf{d}^\top\nabla f(\mathbf{x}^*) \geq c\|\mathbf{d}\| > 0$, we can find an α_0 small enough s.t $\frac{\phi(\alpha)-\phi(0)}{\alpha} > 0 \Rightarrow \phi(\alpha) > \phi(0)$ for sufficiently small α. Hence, $f(\mathbf{x}^*)$ is strict local minimizer.

6.32 Because Ω is convex, let $\phi(\alpha) = f(\mathbf{x}^* + \alpha\mathbf{d})$ for some feasible direction $\mathbf{d}$ and $0 < \alpha_0$ s.t $\mathbf{x}^* + \alpha\mathbf{d} \in \Omega, \forall\alpha \in [0, \alpha_0]$. By Taylor's Theorem, $\phi(\alpha) = f(\mathbf{x}^*) + \alpha\mathbf{d}^\top\nabla f(\mathbf{x}^*) + \frac{\alpha^2}{2}\mathbf{d}^\top F(\mathbf{x}^*)\mathbf{d} + o(\alpha^3)$
$\geq f(\mathbf{x}^*) + \alpha \cdot 0 + \frac{\alpha^2}{2}c\|\mathbf{d}\|^2 + o(\alpha^3) > f(\mathbf{x}^*)$ for sufficiently small α. Hence, $f(\mathbf{x}^*)$ is strict local minimizer.

6.34 Let $\mathbf{u} = [u_1, u_2, \ldots, u_n]^\top$. Using the system
$x_n = \alpha x_{n-1} + \beta u_n$
$= \alpha(\alpha x_{n-1} + \beta u_{n-1}) + \beta u_n$
$\ldots$
$= \beta\sum_{i=1}^{n} \alpha^{n-i}u_i = \mathbf{v}^\top\mathbf{u}$ where $\mathbf{v} = [\beta\alpha^{n-1}, \beta\alpha^{n-1}, \ldots\beta]^\top$
This gives us the quadratic form minimization problem $f(\mathbf{u}) = r\mathbf{u}^\top\mathbf{u} - q\mathbf{v}^\top\mathbf{u}$ with minimizer $\mathbf{u}^* = \frac{q}{2r}\mathbf{v}$ satisfying $\nabla f(\mathbf{u}^*) = 2r\mathbf{u}^* - q\mathbf{v} = 0$ and $F(\mathbf{u}^*) = 2r\mathbf{I}_n > 0$.